

# Principles of Adaptive Array Processing

**Dr. Ulrich Nickel**

Research Institute for High-Frequency Physics and Radar Techniques (FHR)

Research Establishment for Applied Science (FGAN)

53343 Wachtberg

GERMANY

[nickel@fgan.de](mailto:nickel@fgan.de)

## ABSTRACT

*In this lecture we present the principles of adaptive beamforming, the problem of estimating the adaptive weights and several associated practical problems, like preserving low sidelobe patterns, using subarrays and GSCLC configurations. We explain the detection problem with adaptive arrays and the methods for angle estimation. Finally the methods for resolution enhancement (super-resolution methods) are presented.*

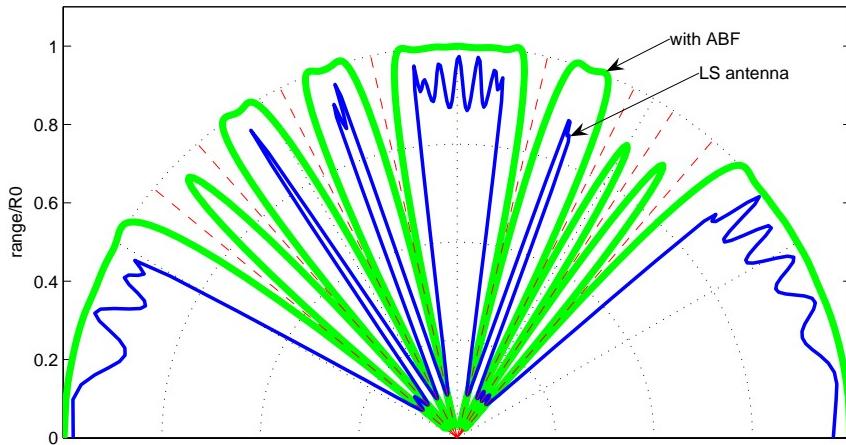
## 1.0 INTRODUCTION – WHY DO WE NEED ADAPTIVE BEAMFORMING

By adaptive beamforming we mean a data dependent modification of the antenna pattern. We are talking here about two types of data: the training data from which we calculate the adaptive weightings and the primary data with which detection and parameter estimation (angle, range, Doppler estimation) is performed. In this section we consider spatial pattern forming only. However, our description is more general to allow the extension to space-time adaptive processing to be developed in the following lectures.

First, one may ask if adaptive beamforming against spatial interference sources is necessary at all, if we can create an antenna with very low sidelobes. A low sidelobe antenna is a fixed spatial filter that does not need training and hence has no adaptation problems. However, the overall reduction of the sidelobes implies an increase of the antenna beamwidth. This effect is not present if single nulls in the directions of the interference are formed. Often the detection and estimation of a target in the vicinity of the interference is of interest (from a jamming point of view co-locating the jammer and the target to be protected is most efficient). Therefore the increased beamwidth would be counter-productive. Figure 1 shows the significant loss of radar visibility with a -40 dB low sidelobe antenna (the 902 element generic array of Lecture 1 with -40 dB Taylor weighting was chosen). Not only is the sidelobe level insufficient to suppress the jamming power, but the broad main beam also makes the antenna blind in the whole sectors with closely spaced jammers. The remarkable property of adaptive antennas is that they can suppress jammers on the skirt of the main beam such that the radar can look through between two neighbouring jammers.

Adaptive antennas can only suppress interference that was measured in the training data. To be protected against surprise interference or insufficiently measured interference a compromise solution using a moderate low sidelobe antenna with simultaneous adaptive beamforming is often sought.

<b>Report Documentation Page</b>			Form Approved OMB No. 0704-0188	
<p>Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p>				
1. REPORT DATE <b>01 SEP 2006</b>	2. REPORT TYPE <b>N/A</b>	3. DATES COVERED <b>-</b>		
4. TITLE AND SUBTITLE <b>Principles of Adaptive Array Processing</b>			5a. CONTRACT NUMBER	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)			5d. PROJECT NUMBER	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Research Institute for High-Frequency Physics and Radar Techniques (FHR) Research Establishment for Applied Science (FGAN) 53343 Wachtberg GERMANY</b>			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release, distribution unlimited</b>				
13. SUPPLEMENTARY NOTES <b>See also ADM001925, Advanced Radar Signal and Data Processing., The original document contains color images.</b>				
14. ABSTRACT				
15. SUBJECT TERMS				
16. SECURITY CLASSIFICATION OF:  a. REPORT      b. ABSTRACT      c. THIS PAGE <b>unclassified</b> <b>unclassified</b> <b>unclassified</b>			17. LIMITATION OF ABSTRACT <b>UU</b>	18. NUMBER OF PAGES <b>20</b>
19a. NAME OF RESPONSIBLE PERSON				



**Figure 1: Range reduction with 10 jammers: Low sidelobe antenna vs. ABF Antenna with 902 elements and -40dB Taylor weighting, ABF with 32 subarrays.**

## 2.0 OPTIMUM ADAPTIVE BEAMFORMING

Suppose we know the interference situation, i.e. if we know the interference covariance matrix  $\mathbf{Q}$ , what is the optimum beamforming vector  $\mathbf{w}$ ? From the Likelihood Ratio test criterion we know that the probability of detection is maximised if we choose the weight vector that maximises the signal-to-noise-plus-interference ratio (SNIR) for a given (expected) signal  $\mathbf{a}_0$ ,

$$\max_{\mathbf{w}} \left| \mathbf{w}^H \mathbf{a}_0 \right|^2 / E \left\{ \left| \mathbf{w}^H \mathbf{n} \right|^2 \right\} = \max_{\mathbf{w}} \left| \mathbf{w}^H \mathbf{a}_0 \right|^2 / (\mathbf{w}^H \mathbf{Q} \mathbf{w}). \quad (1)$$

The solution of this optimisation is

$$\mathbf{w} = \mu \mathbf{Q}^{-1} \mathbf{a}_0 \quad \text{with} \quad \mathbf{Q} = E \left\{ \mathbf{n} \mathbf{n}^H \right\}. \quad (2)$$

$\mu$  is a free normalisation constant.

Sometimes only the jamming power is minimised under additional constraints, e.g.  $\mathbf{w}^H \mathbf{c}_i = k_i$ , for suitable vectors  $\mathbf{c}_i$  and numbers  $k_i$ ,  $i = 1..r$ . Although this is an intuitively reasonable criterion it does not necessarily give the maximum SNIR and hence maximum probability of detection. For certain constraints both solutions are however equivalent. The constrained optimisation problem can be written in general terms as

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{Q} \mathbf{w} \quad \text{s.t.} \quad \mathbf{w}^H \mathbf{C} = \mathbf{k} \quad \left( \mathbf{w}^H \mathbf{c}_i = k_i, i = 1..r \right) \quad (3)$$

and the solution can be written as

$$\mathbf{w} = \sum_{i=1}^r \lambda_i \mathbf{Q}^{-1} \mathbf{c}_i = \mathbf{Q}^{-1} \mathbf{C} \left( \mathbf{C}^H \mathbf{Q}^{-1} \mathbf{C} \right)^{-1} \mathbf{k}. \quad (4)$$

Special cases:

1. Single unit gain constraint:  $\mathbf{w}^H \mathbf{a}_0 = 1 \Rightarrow \mathbf{w} = (\mathbf{a}_0^H \mathbf{Q}^{-1} \mathbf{a}_0)^{-1} \mathbf{Q}^{-1} \mathbf{a}_0$ . This is obviously equivalent to the SNIR-optimum solution (2).

2. Gain and derivative constraint:  $\mathbf{w}^H \mathbf{a}_0 = 1, \quad \mathbf{w}^H \mathbf{a}'_0 = 0 \Rightarrow \mathbf{w} = \mu \mathbf{Q}^{-1} \mathbf{a}_0 + \kappa \mathbf{Q}^{-1} \mathbf{a}'_0$ . A derivative constraint is added to make the weight less sensitive against mismatch of the steering direction.
3. Gain and norm constraint:  $\mathbf{w}^H \mathbf{a}_0 = 1, \quad \mathbf{w}^H \mathbf{w} = c \Rightarrow \mathbf{w} = \mu (\mathbf{Q} + \delta \mathbf{I})^{-1} \mathbf{a}_0$ . The norm constraint is added to make the weight numerically stable. This is equivalent to the famous diagonal loading technique which we will consider later.
4. Norm constraint only:  $\mathbf{w}^H \mathbf{w} = 1 \Rightarrow \mathbf{w} = \min EV(\mathbf{Q})$ . Without a gain constraint the weight vector produces a nearly omni-directional pattern, but with nulls in the interference directions. This is also called the power inversion weight, because the pattern displays the inverted interference power.

As we mentioned before, the fulfilment of the constraints may imply a significant loss in SNIR. Therefore one sometimes aims at a compromise between power minimisation and the constraints which can be introduced by coupling factors  $b_i$ . This is called the soft constraint optimisation

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{Q} \mathbf{w} + \sum_{i=1}^r b_i |\mathbf{w}^H \mathbf{c}_i - k_i|^2 \quad \text{or} \quad \min_{\mathbf{w}} \mathbf{w}^H \mathbf{Q} \mathbf{w} + (\mathbf{w}^H \mathbf{C} - \mathbf{k})^H \mathbf{B} (\mathbf{w}^H \mathbf{C} - \mathbf{k}) \quad (5)$$

with  $\mathbf{B} = \text{diag}\{b_1, \dots, b_r\}$ . The solution of the soft-constraint optimisation is

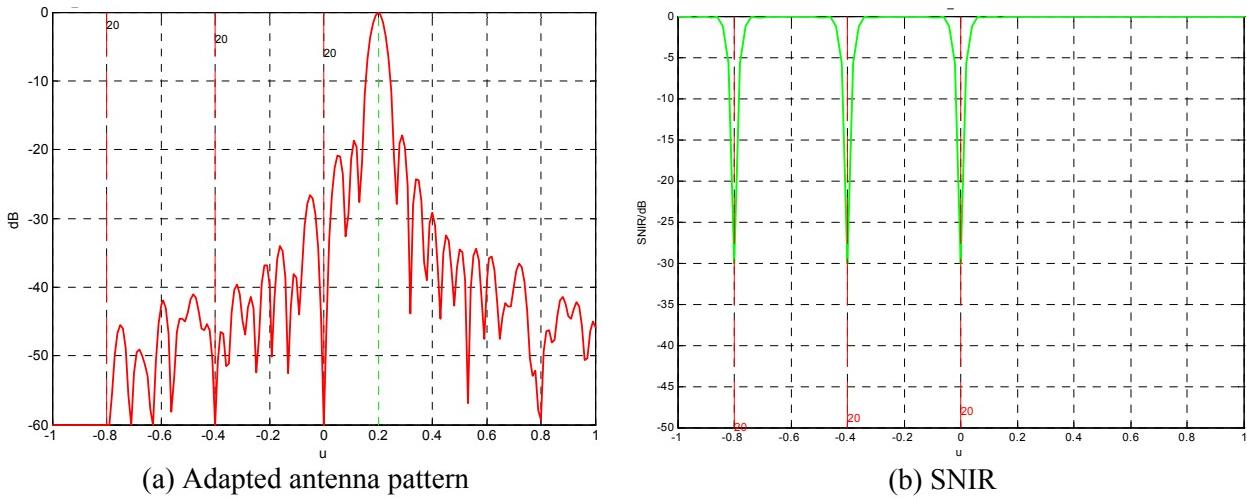
$$\mathbf{w} = (\mathbf{Q} + \mathbf{C} \mathbf{B} \mathbf{C}^H)^{-1} \mathbf{C} \mathbf{B} \mathbf{k}. \quad (6)$$

Typical special cases are:

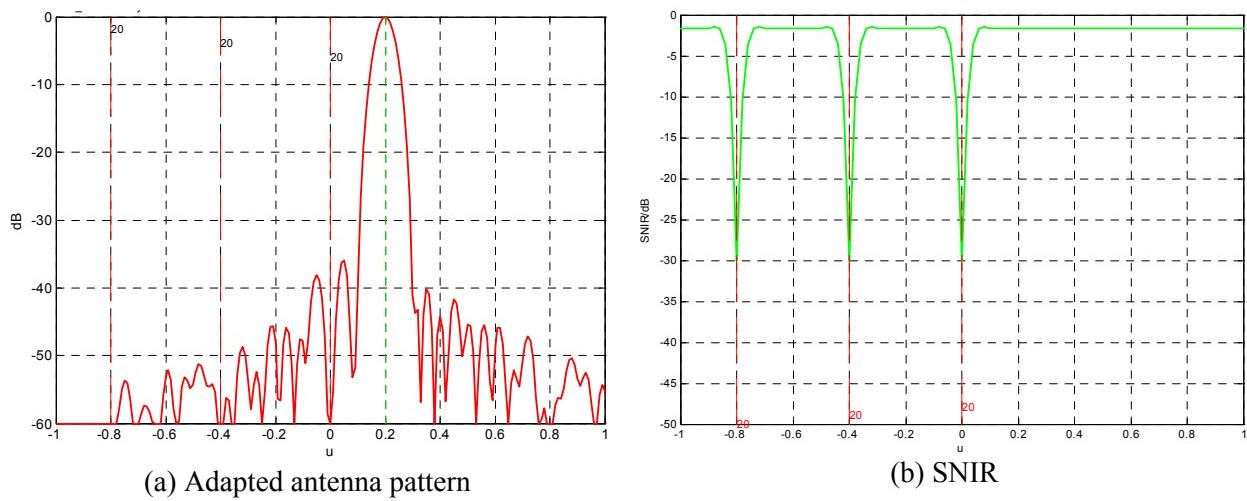
1. Single unit gain constraint:  $\mathbf{w}^H \mathbf{a}_0 = 1 \Rightarrow \mathbf{w} = \mu (\mathbf{Q} + b \mathbf{a}_0 \mathbf{a}_0^H)^{-1} \mathbf{a}_0$ . This corresponds to an inclusion of the desired signal into the interference covariance matrix. As we will show later this inclusion has no effect on the SNIR.
2. Gain and derivative constraint:  $\mathbf{w}^H \mathbf{a}_0 = 1, \quad \mathbf{w}^H \mathbf{a}'_0 = 0 \Rightarrow \mathbf{w} = \mu (\mathbf{Q} + b_1 \mathbf{a}_0 \mathbf{a}_0^H + b_2 \mathbf{a}'_0 \mathbf{a}'_0^H)^{-1} \mathbf{a}_0$ .

The performance of ABF is often displayed by the adapted antenna pattern. A typical adapted antenna pattern with 3 jammers of 20 dB SNR is shown in Figure 2(a). This pattern does however not show how the actual jamming power and the null depth play together. Plots of the SNIR are better suited for displaying this effect. The SNIR is typically plotted for varying target directions while the interference scenario is held fixed, as seen in Figure 2(b). The SNIR is normalised to the SNR in absence of any jamming and without ABF. In other words, this pattern shows the insertion loss arising from the jamming scenario with applied ABF. The effect of target and steering direction mismatch is not accounted for in the SNIR plot. This effect is displayed by the scan pattern, i.e. the pattern arising if the adapted array scans over a fixed target and interference scenario. Such a plot is rarely shown because of the many parameters to be varied.

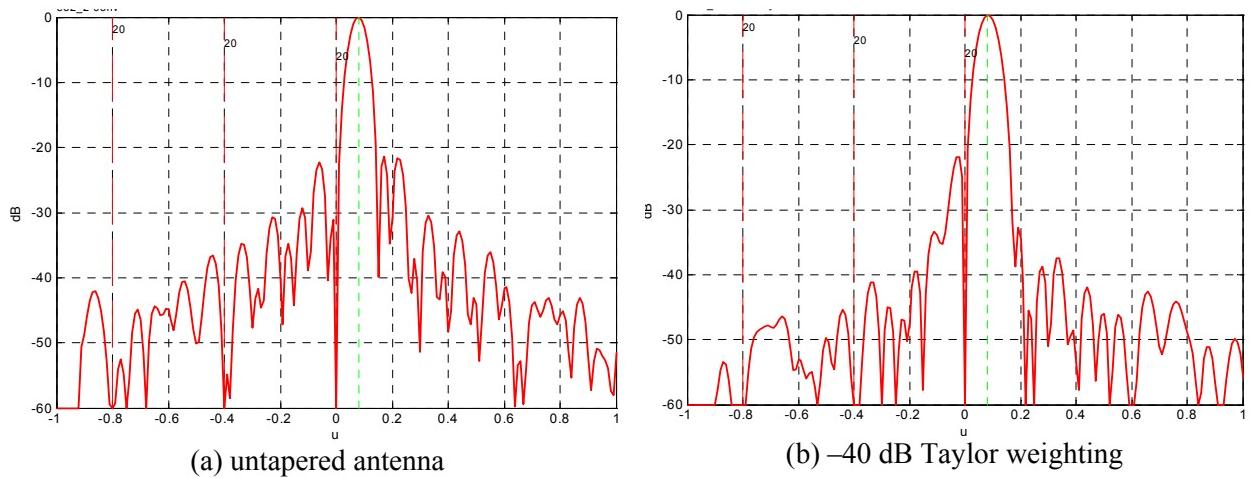
Figure 2 shows the case of an untapered planar antenna. The first sidelobes of the unadapted antenna pattern are at -17 dB and are nearly unaffected by the adaptation process. If we have an antenna with low sidelobes, the peak sidelobe level is much more affected. This is seen in Figure 3. Due to the tapering we have a loss in SNIR of 1.7 dB compared to the reference antenna (untapered without ABF and jamming).



**Figure 2: Antenna and normalised SNIR patterns for a three jammer configuration.**



**Figure 3: Antenna and normalised SNIR patterns for a three jammer configuration for antenna with low sidelobes (-40 dB Taylor weighting).**



**Figure 4: Antenna patterns for mainbeam jamming configuration.**

The low-sidelobe antenna is also much more sensitive against mainbeam jamming. This is mainly due to the broader mainbeam width. In summary, if we want to look through between closely spaced jammers i.e. if mainbeam jamming is an issue, the low-sidelobe tapering is counter-productive. Low sidelobes are only appropriate against completely unknown extended interference. Any a-priori knowledge on the extent and direction of the interference should otherwise be better countered by some deterministic pattern shaping as presented in Lecture 1.

### 3.0 ESTIMATION OF ADAPTIVE WEIGHTS

In reality the interference covariance matrix is not known and must be estimated from some training data  $\mathbf{Z} = (\mathbf{z}_1, \dots, \mathbf{z}_K)$ . The maximum likelihood estimate of the covariance matrix is

$$\hat{\mathbf{Q}}_{SMI} = \frac{1}{K} \sum_{k=1}^K \mathbf{z}_k \mathbf{z}_k^H. \quad (7)$$

This is called the Sample Matrix Inversion algorithm (SMI). The SMI method is only asymptotically a good estimate. For small sample size it is known to be not very stable. For the matrix being invertible we need at least  $K = N$  samples. According to Brennan's Rule, [1], one needs  $2K$  samples to obtain on the average a SNIR loss below 3 dB. For smaller sample size the performance can be considerably worse. However, by simply adding a multiple of the identity matrix to the estimated covariance matrix a nearly optimum performance can be achieved. This is called the loaded sample matrix estimate (LSMI).

$$\hat{\mathbf{Q}}_{LSMI} = \frac{1}{K} \sum_{k=1}^K \mathbf{z}_k \mathbf{z}_k^H + \delta \cdot \mathbf{I} \quad (8)$$

The difference between SMI and LSMI is drastically shown in Figure 5. For a “reasonable” choice of the loading factor (a rule of thumb is  $\delta = 2\sigma^2 \dots 4\sigma^2$  for an untapered antenna) we need only  $2M$  snapshots to obtain an average SNIR loss below 3 dB, if  $M$  denotes the number of jammers (dominant eigenvalues) present, [3]. So the sample size can be considerably below the dimension of the matrix. The effect of the loading factor is that the dynamic range of the small eigenvalues is compressed. The small eigenvalues possess the largest statistical fluctuation, but have the greatest influence on the weight fluctuation after inversion of the matrix.

One may go even further and ignore the small eigenvalue estimates completely, i.e. one tries to find an estimate of the inverse covariance matrix based on the dominant eigenvectors and eigenvalues. For high SNR we can replace the inverse covariance matrix by a projection matrix. Suppose we have  $M$  jammers with amplitudes  $b_1(t), \dots, b_M(t)$  in directions  $\mathbf{u}_1, \dots, \mathbf{u}_M$ . The received data can then be written as  $\mathbf{z}(t_k) = \sum_{m=1}^M \mathbf{a}(\mathbf{u}_m) b_m(t_k) + \mathbf{n}(t_k)$ , or short in vector-matrix notation  $\mathbf{z}(t_k) = \mathbf{Ab}(t_k) + \mathbf{n}(t_k)$ . Then it is easy to show that

$$\mathbf{Q} = \mathbf{ABA}^H + \mathbf{I}, \quad (9)$$

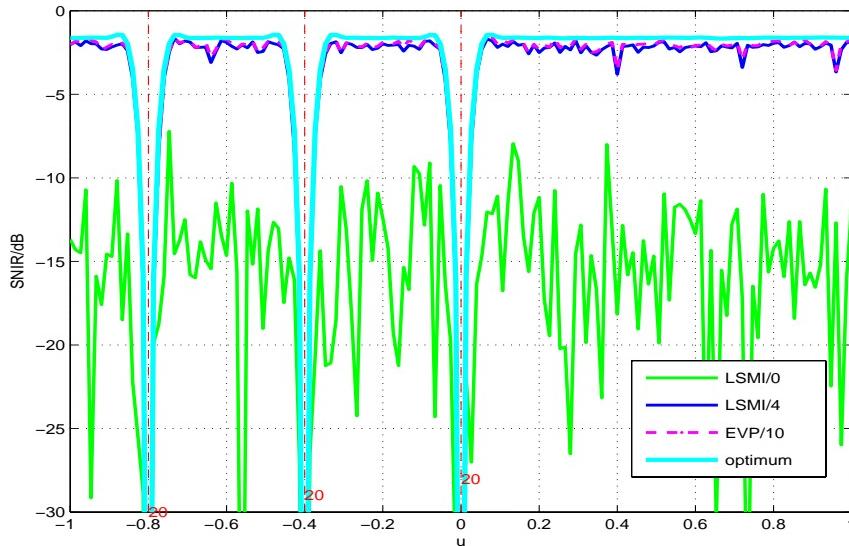
where we have normalised the noise power to 1 and  $\mathbf{B} = E\{\mathbf{bb}^H\}$ . Using the matrix inversion lemma we have now for the inverse that

$$\mathbf{Q}^{-1} = \mathbf{I} - \mathbf{A}(\mathbf{B}^{-1} + \mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \xrightarrow{\mathbf{B} \rightarrow \infty} \mathbf{I} - \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H = \mathbf{P}_{\mathbf{A}}^{\perp}. \quad (10)$$

$\mathbf{P}_{\mathbf{A}}^{\perp}$  is a projection on the space orthogonal to the columns of  $\mathbf{A}$ . For strong jammers we may therefore replace the estimated inverse covariance matrix by a projection built from the dominant eigenvectors of

the estimated covariance matrix, called EVP method. As the eigenvectors  $\mathbf{X}$  are usually orthonormalised the projection can be written as  $\mathbf{P}_X^\perp = \mathbf{I} - \mathbf{XX}^H$ .

Figure 5 shows the performance of SMI, LSMI and EVP for a 32 channel array and 32 data snapshots (the generic planar array of Lecture 1) and three jammers of 20 dB input JNR. Note the drastic improvement by LSMI and EVP and the little difference between LSMI and EVP. The three methods are based on the same realisation of covariance estimate. What is not visible in this figure is the sidelobe level. EVP can better preserve a desired low sidelobe level than LSMI.



**Figure 5: SNIR for SMI, LSMI ( $\delta = 4\sigma^2$ ) and eigenvector projection (dim(Jammer subspace)=3).**

For EVP we have to know the dimension of the jammer subspace (dimJSS). In complicated scenarios and with channels errors this can be difficult to determine. If dimJSS is grossly overestimated a loss in SNIR occurs, if dimJSS is underestimated the jammers are not fully suppressed. One is therefore interested in subspace methods with less sensitivity against the choice of subspace dimension. This property is achieved by a “weighted projection”, i.e. by replacing the projection by a matrix

$$\mathbf{P}_{LMI} = \mathbf{I} - \mathbf{X}\mathbf{D}\mathbf{X}^H, \quad (11)$$

where  $\mathbf{X}$  is a set of orthonormal vectors spanning the interference subspace and  $\mathbf{D}$  is a diagonal weighting matrix.  $\mathbf{P}_{LMI}$  is of course no projection. This type of methods are called lean matrix inversion methods (LMI).

Different kinds of adaptive weight estimation methods have been proposed which can all be interpreted as an LMI method (e.g. constrained signal subspace method, CSSP, constrained noise subspace method CNSP, [2]). The LMI matrix can also be economically calculated by an eigenvector-free QR-decomposition method, [3].

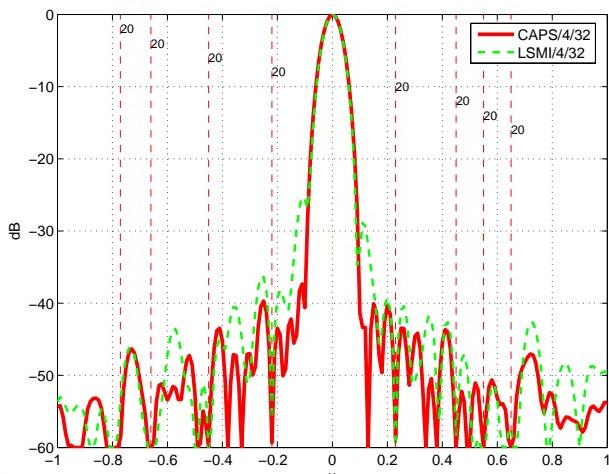
One of the most efficient methods for pattern stabilisation while maintaining a low desired sidelobe level is the constrained adaptive pattern synthesis (CAPS) algorithm, [4]. Suppose we have a weight vector for beamforming with low sidelobes in a certain direction,  $\mathbf{m}$ . In full generality the CAPS weight can be written as

$$\mathbf{w}_{CAPS} = \frac{1}{\mathbf{m}^H \hat{\mathbf{R}}^{-1} \mathbf{m}} \hat{\mathbf{R}}^{-1} \mathbf{m} - \mathbf{X}_\perp \left( \mathbf{X}_\perp^H \mathbf{C} \mathbf{X}_\perp \right)^{-1} \mathbf{X}_\perp^H \mathbf{C} \left( \frac{1}{\mathbf{m}^H \hat{\mathbf{R}}^{-1} \mathbf{m}} \hat{\mathbf{R}}^{-1} \mathbf{m} - \mathbf{m} \right), \quad (12)$$

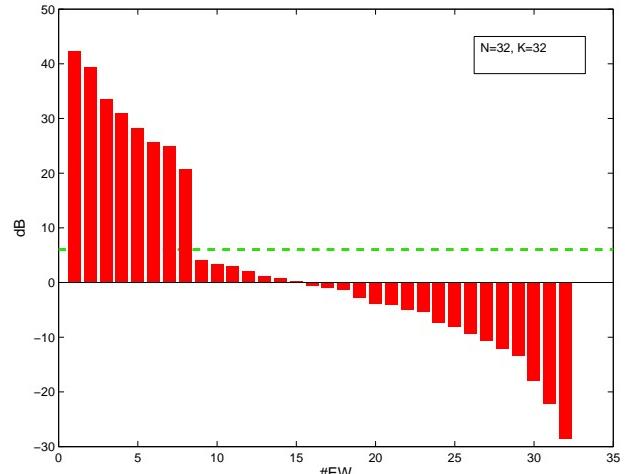
where the columns of the matrix  $\mathbf{X}_\perp$  span the space orthogonal to  $[\mathbf{X}, \mathbf{m}]$  and  $\mathbf{X}$  is again a unitary  $L \times M$  matrix with columns spanning the interference subspace which is assumed to be of dimension  $M$ .  $\mathbf{C}$  is a directional weighting matrix,  $\mathbf{C} = \int_{\Omega} \mathbf{a}(\mathbf{u}) \mathbf{a}(\mathbf{u})^H p(\mathbf{u}) d\mathbf{u}$ , if  $\Omega$  denotes the set of directions of interest and  $p(\mathbf{u})$  a directional weighting function. If we use no directional weighting,  $\mathbf{C} \approx \mathbf{I}$ , the CAPS weight vector simplifies to

$$\mathbf{w}_{CAPS} = \mathbf{m} + \mathbf{P}_{[\mathbf{X}, \mathbf{m}]} \left( \frac{1}{\mathbf{m}^H \hat{\mathbf{R}}^{-1} \mathbf{m}} \hat{\mathbf{R}}^{-1} \mathbf{m} - \mathbf{m} \right), \quad (1.13)$$

where  $\mathbf{P}_{[\mathbf{X}, \mathbf{m}]}$  denotes the projection onto the space spanned by the columns of  $\mathbf{X}$  and  $\mathbf{m}$ . Only for very low sidelobe antennas (e.g. -40 dB) and severe jamming situations (e.g. mainbeam jamming) a difference to LSMI and LMI is visible, see Figure 6(a) for the scenario of Figure 1. The loading factor was set to  $\delta = 4\sigma^2$ .



(a) Adapted patterns with LSMI and CAPS



(b) Eigenvalues of 10 jammer scenario

**Figure 6: LSMI and CAPS for scenario with 10 jammers, generic array with 32 subarrays, K=32.**

All subspace methods require an estimate of the dimension of the interference subspace ( $\text{dimJSS}$ ). Usually this is derived from the eigenvectors. For complicated scenarios and small sample size a clear decision what constitutes a dominant eigenvalue may be difficult. Figure 6(b) shows the eigenvalues for the 10 jammer scenario estimated from 32 data snapshots (the loading factor level for LSMI is indicated by the green dashed line). However, this is a simple scenario with equally strong jammers and without channel errors. Several criteria to determine  $\text{dimJSS}$  have been suggested: the Akaike Information Criterion (AIC), Minimum Description Length (MDL), and threshold tests, see e.g. [5]. The AIC and MDL are based on the test criterion

$$T(m) = \frac{\frac{1}{N-m} \sum_{i=m+1}^N \lambda_i}{\left( \prod_{i=m+1}^N \lambda_i \right)^{1/(N-m)}}, \quad (14)$$

where  $\lambda_i$  denote the eigenvalues of the estimated covariance matrix ordered in decreasing magnitude. For AIC resp. MDL one chooses  $\text{dimJSS}$  as the argument  $\hat{M}$  where the following functions are minimum

$$\begin{aligned} AIC(m) &= K(N-m) \log[T(m)] + m(2N-m) \\ MDL(m) &= K(N-m) \log[T(m)] + (m/2)(2N-m) \log K, \end{aligned} \quad (15)$$

The typical threshold test for known noise power  $\sigma^2$  uses the statistic  $L(m)$  based on the ordered eigenvalues

$$L(m) = \frac{2K}{\sigma^2} \sum_{i=m+1}^N \lambda_i \quad (16)$$

and the decision is found if the *test statistic is for the first time below the threshold*

```

for i=1..N do
    if  $L(m) \leq \chi^2_{2K(N-m),\alpha}$  :  $\hat{M} = m$ ; STOP;
    end

```

The symbol  $\chi^2_{r,\alpha}$  denotes the  $\alpha$ -percentage point of the  $\chi^2$ -distribution with  $r$  degrees of freedom. The probability to overestimate dimJSS is then asymptotically bounded by  $\alpha$ . As explained in [5] diagonal loading can improve AIC and MDL for small sample size and make these criteria robust against errors.

The eigenvector-free weight calculation methods use different criteria based on QR-decomposition, [3].

## 4.0 PROBLEMS OF ADAPTIVE BEAMFORMING

In the preceding section we have already mentioned some of the problems associated with adaptive arrays. We give here a full list of problems and the possible mitigations.

### 4.1 Numerical Inversion of the Matrix

This should be no more a problem with today's floating point processors. For training data  $\mathbf{Z} = (\mathbf{z}_1, \dots, \mathbf{z}_K)$  the estimated covariance matrix  $\hat{\mathbf{Q}} = \frac{1}{K} \mathbf{Z} \mathbf{Z}^H$  is positive definite and hermitean. Cholesky-decomposition is therefore suitable and numerically stable. For solving the linear equation we can use forward-backward recursion for the triangular matrices. For small sample size also QR-decomposition of the data matrix  $\mathbf{Z}^H$  is suitable, because for  $\mathbf{Z}^H = \mathbf{U} \mathbf{R}$  with  $\mathbf{U}$  unitary and  $\mathbf{R}$  upper triangular,  $\hat{\mathbf{Q}} = \frac{1}{K} \mathbf{R}^H \mathbf{U}^H \mathbf{U} \mathbf{R} = \frac{1}{K} \mathbf{R}^H \mathbf{R}$ . This means that up to a normalisation constant the upper triangular matrix  $\mathbf{R}$  is the same as the Cholesky factor of  $\hat{\mathbf{Q}}$ . There are also efficient recursive QR algorithms to update a given  $\mathbf{R}_K$  if a new data snapshot  $\mathbf{z}_{K+1}$  comes in, [6].

### 4.2 Low Number of Data Samples

Diagonal loading and subspace methods are the appropriate means to counter these effects.

### 4.3 Weight Fluctuation

The weight fluctuation is basically a consequence of a too low number of data samples. However, depending on the required sidelobe level "too low" may mean different numbers. Diagonal loading and subspace methods are the appropriate means to counter these effects.

#### 4.4 Preservation of Low Sidelobes

The first point to note here is that, if the quiescent channels (only receiver noise is present) have unequal power, the adaptive processor will try to make the power equal, such that the resulting covariance matrix of these channels is the identity matrix. This means that all tapering applied before adaptation will be retracted by the adaptive processor. To avoid this we have to normalise all channels, such that the noise power is equal. All taper weights have to be accounted for in the desired (or model) signal appearing in the constraints ( $\mathbf{a}_0$  in (2), (4), (6) or  $\mathbf{m}$  in (12)). This is of particular importance if we work with subarrays, see the following section.

As mentioned before, additional constraints, diagonal loading, subspace methods, and in particular the CAPS method are suited to preserve the low sidelobes.

#### 4.5 Real Time Requirement (reduction of the number of channels)

Much work has been done on fast weight estimation. However, if we count the number of operations the application of the weights to the data, i.e. the digital beamforming operation, gives the largest computational load, because we have in general a large number of range cells. Therefore there is an interest in reducing the number of channels. In the early days of adaptive antennas one simply added some auxiliary elements by which the interference was estimated and then subtracted from the main antenna. This is called the sidelobe canceller (SLC), see Figure 7. The main antenna can be even a conventional dish antenna.

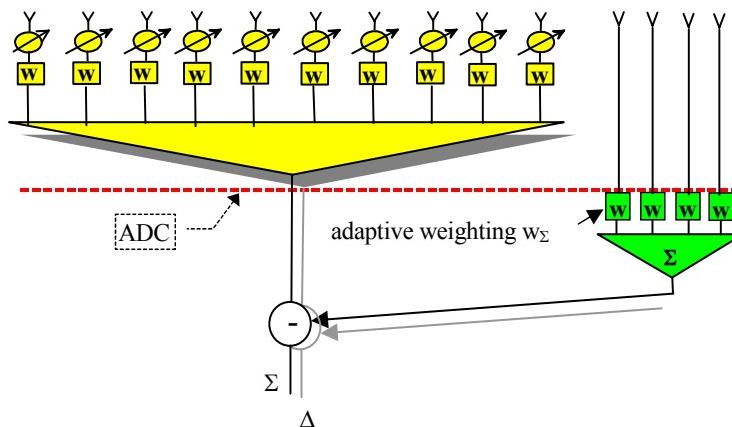


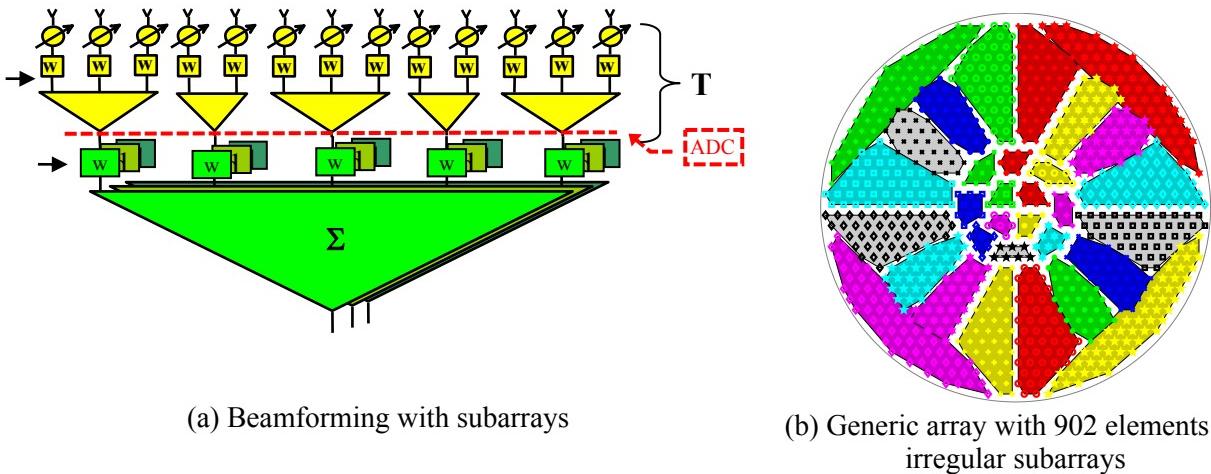
Figure 7: Principle of sidelobe canceller.

The adaptive weights  $\mathbf{w}$  are determined such that the jamming power estimated with the auxiliary channel  $\mathbf{x}$  is optimally subtracted from the main channel  $y$  in a mean square sense:  $E\{|y - \mathbf{w}^H \mathbf{x}|^2\} = \min!$ . The solution of this optimisation is  $\mathbf{w}_{slc} = E\{\mathbf{x}\mathbf{x}^H\}^{-1} \cdot E\{\mathbf{x}y\} = \mathbf{Q}_{xx}^{-1} \mathbf{q}_{xy}$ . One can show that this solution is equal to the SNR-optimising solution of (2)  $\mathbf{w} = \mu \mathbf{Q}^{-1} \mathbf{a}_0$ , if we set the beamforming vector  $\mathbf{a}_0 = (1, 0, \dots, 0)^T$  and partition the covariance matrix into the main and auxiliary channel part  $\mathbf{Q} = \begin{pmatrix} r & \mathbf{q}_{xy} \\ \mathbf{q}_{xy}^H & \mathbf{Q}_{xx} \end{pmatrix}$ . The beamforming (signal model) vector shows that this method is SNR-optimum, if the main antenna has a high gain such the signal in the auxiliary channel can be neglected.

For phased arrays forming subarrays is a suitable means of reducing the number of channels. We can apply the SNR-optimum weight to the subarray outputs. However, a fully filled array on a regular grid

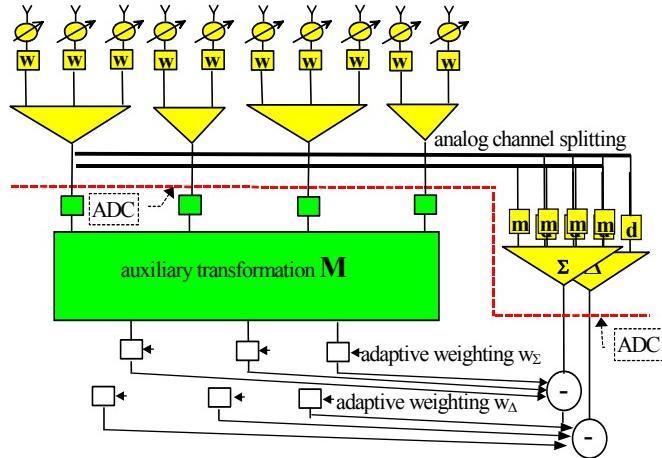
grouped into equal subarrays will produce grating notches in adaptive beamforming, because the subarray centres are separated by more than  $\lambda/2$ . Grating effects can be avoided by choosing subarrays of different size and irregular shape as in the generic subarray of Lecture 1, Figure 8. The problem is that now the receiver noise powers at each subarray are unequal (assuming that the amplifiers of the TRMs create the main noise contribution). The adaptation process will try to whiten the noise which means it will try to remove the low sidelobe taper as much as possible on subarray level. This can be avoided by normalising to equal noise power as mentioned in Lecture 1. If the whole subarray forming process is described by a matrix  $\mathbf{T}$  (including phase shifting and full tapering, Lecture 1), then the normalisation at the subarray outputs should be such that  $\mathbf{T}^H \mathbf{T} = \mathbf{I}$ . If the whole conventional beamforming is then done by  $\mathbf{w} = \mathbf{T} \mathbf{1}$ , where  $\mathbf{1}$  is the vector of ones (of length equal to no of subarrays), then we decompose  $\mathbf{T} = \mathbf{T}_0 \mathbf{D}$ , such that  $\mathbf{T}_0^H \mathbf{T}_0 = \mathbf{I}$  and use as beamforming vector at the subarray outputs  $\mathbf{m} = \mathbf{D} \mathbf{1}$ . The covariance matrix is calculated from the subarray output data as usual, e.g.

$$\hat{\mathbf{Q}} = \frac{1}{K} \sum_{k=1}^K \mathbf{T}^H \mathbf{z}_k \mathbf{z}_k^H \mathbf{T} = \mathbf{T}^H \frac{1}{K} \mathbf{Z} \mathbf{Z}^H \mathbf{T} = \mathbf{T} \hat{\mathbf{Q}} \mathbf{T} .$$



**Figure 8: Beamforming with subarrays.**

Subarrays and the SLC concept can be combined to a Generalised Sidelobe Canceller (GSLC) as indicated in Figure 9. In this case one can apply a matrix operation  $\mathbf{M}$  to the subarray outputs. If this matrix is orthogonal to the main beam beamforming vector  $\mathbf{m}$ , i.e. if  $\mathbf{M}^H \mathbf{m} = \mathbf{0}$ , the so called signal blocking condition is fulfilled, then this GSLC is mathematically identical to the SNR-optimum weight (2). The advantage of the GSLC can be that the main channels may be formed in an analogue manner, resulting in a reduced sensitivity to AD-converter limiting for strong jammers in the main channel, because the jammers are attenuated by the sidelobes, and in a possibly reduced sensitivity to bandwidth errors as the analogue beamforming network has a broadband characteristic.

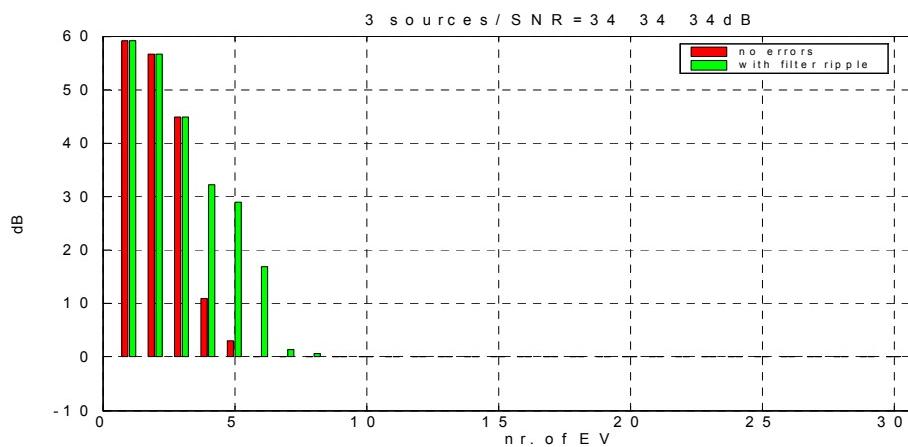


**Figure 9: Generalised Sidelobe Canceller configuration.**

#### 4.6 Sensitivity Against Errors

A big advantage of adaptive beamforming versus deterministic beamforming is that for interference cancellation the steering vectors into the interference directions need not to be known. Array element position errors therefore have no influence on the interference cancellation (they do have an influence on the beam shape into the target direction). The error bounds derived in Lecture 1 Sect. 4.4 on weight errors can be applied here, too.

A key issue in ABF is the equality of the channels. All fixed errors (amplitude, phase) are not relevant for the suppression, because the covariance matrix measures only the correlation between the channels. Fixed offsets have the effect of an additional jammer. However, random errors which lead to a decorrelation have an impact. Typically these are I- and Q-demodulation errors (different amplification, orthogonality errors) and differences in the bandpass filter characteristic. These error effects have been analysed in [7]. The effect of these errors is a leakage of the dominant eigenvalues into the noise eigenvalues. With more degrees of freedom such errors can be partly compensated. Figure 10 shows the eigenvalues for a scenario with 3 jammers with the generic array (32 subarrays) with and without bandpass filter errors (1 dB amplitude ripple). The apparent number of jammers has been doubled by the errors.



**Figure 10: Eigenvalue leakage due to bandpass filter ripple.**

## 4.7 Inclusion of the Signal in the Training Data

The data for estimating the adaptive filter weights (training data) should not contain the signal, because it would be considered as interference and suppressed. Only if the signal present is exactly equal to the expected signal (the beamforming vector), then signal inclusion does not have an effect, because for  $\mathbf{R} = \mathbf{s}_0 \mathbf{s}_0^H + \mathbf{Q}$  we have  $\mathbf{R}^{-1} = \mathbf{Q}^{-1} - \frac{\mathbf{Q}^{-1} \mathbf{s}_0^H \mathbf{Q}^{-1}}{1 + \mathbf{s}_0^H \mathbf{Q}^{-1} \mathbf{s}_0}$  such that  $\mathbf{R}^{-1} \mathbf{a}_0 = \mathbf{Q}^{-1} \mathbf{a}_0 - \frac{\mathbf{s}_0^H \mathbf{Q}^{-1} \mathbf{a}_0}{1 + \mathbf{s}_0^H \mathbf{Q}^{-1} \mathbf{s}_0} \mathbf{Q}^{-1} \mathbf{s}_0$ . Hence, if  $\mathbf{a}_0 \parallel \mathbf{s}_0$  we have  $\mathbf{Q}^{-1} \mathbf{a}_0 \parallel \mathbf{R}^{-1} \mathbf{a}_0$ .

We have to take care that the training data are taken where no target is expected (dead zones) or we have to exclude the cell under test from the training data. If the signal is weak (e.g. if the transmit pulse is a long coded pulse) then signal cancellation will have little effect.

## 4.8 Non-Stationary Interference

The general adaptation scheme consists of the training phase followed by a weight application phase with weights frozen as indicated in the figure below. This method is susceptible to any kind of non-stationarity of the jamming scenario.

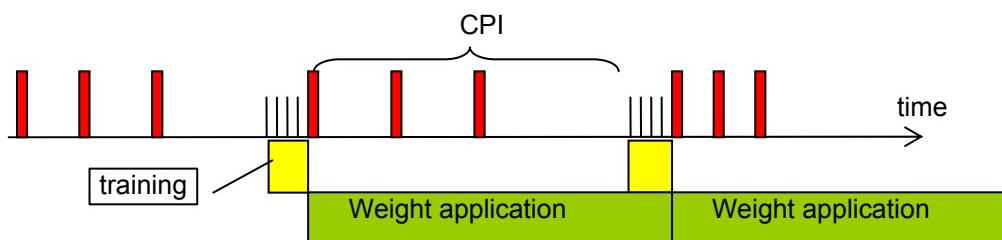


Figure 11: Adaptation by training and weight freezing.

If the weights are applied in the same time window where the training data have been taken, no time delay occurs. We have only introduced some latency of the outputs. Moreover, if the jammer moves in the training phase, an extended null will be created and thus a reasonable suppression can be achieved. It is also possible to force the creation of extended nulls to account for moving interference. This has been done in [8]. Another option against moving interference is to calculate time-varying adaptive weights by a linear extrapolation. This has been presented in [9].

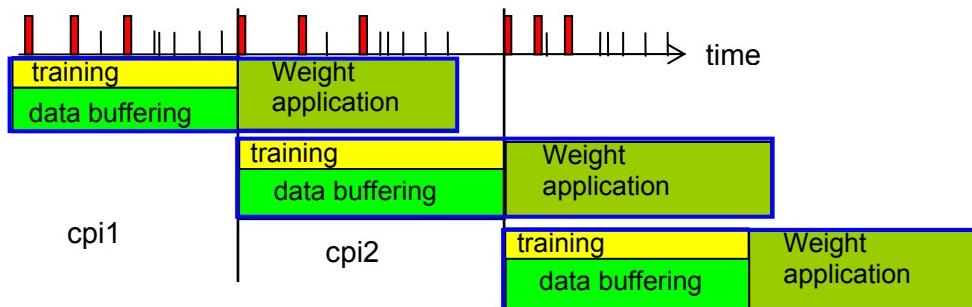


Figure 12: Training and weight application on the same data.

## 5.0 ADAPTIVE DETECTION

In Section 2.0 we have only considered the maximisation of the signal-to-noise ratio, which is equivalent to maximising the detection probability for the case of known interference covariance. This can be derived

from the Likelihood ratio test presented in Lecture 1. In reality the interference covariance matrix is estimated and the correct maximum-likelihood test must consider both, the secondary (training) and primary (weight application) data sets.

The likelihood ratio test for the adaptive beamforming problem with primary and secondary data has been studied in [10] and is often known as the Generalised Likelihood Ratio Test (GLRT), although this term describes a class of test procedures independent of ABF. The resulting test has the structure

$$\text{Decide for } \begin{cases} \mathcal{H} \text{ (no target present)} \\ \mathcal{A} \text{ (target present)} \end{cases} \quad \text{if} \quad \begin{cases} T_{GLRT}(\mathbf{z}) \leq \eta \\ T_{GLRT}(\mathbf{z}) > \eta \end{cases} \quad (18)$$

with

$$T_{GLRT}(\mathbf{z}) = \frac{\left| \mathbf{a}_0^H \hat{\mathbf{Q}}^{-1} \mathbf{z} \right|^2}{\mathbf{a}_0^H \hat{\mathbf{Q}}^{-1} \mathbf{a}_0 \left( 1 + \frac{1}{K} \mathbf{z}^H \hat{\mathbf{Q}}^{-1} \mathbf{z} \right)}, \quad (19)$$

where  $\mathbf{a}_0$  is the plane wave response (the expected signal),  $\mathbf{z}$  is the primary data vector, and  $\hat{\mathbf{Q}} = \frac{1}{K} \sum_{k=1}^K \mathbf{z}_k \mathbf{z}_k^H$  is the covariance matrix estimated from the secondary data.

A test based only on the adaptive beam output has also been suggested, called the Adaptive Matched Filter (AMF) test, [12],

$$T_{AMF}(\mathbf{z}) = \frac{\left| \mathbf{a}_0^H \hat{\mathbf{Q}}^{-1} \mathbf{z} \right|^2}{\mathbf{a}_0^H \hat{\mathbf{Q}}^{-1} \mathbf{a}_0}. \quad (20)$$

By means of different considerations using subspaces and invariances a detector called Adaptive Cosine Estimator (ACE), [13], has been introduced, which is closely related to the detectors mentioned before.

$$T_{ACE}(\mathbf{z}) = \frac{\left| \mathbf{a}_0^H \hat{\mathbf{Q}}^{-1} \mathbf{z} \right|^2}{\mathbf{a}_0^H \hat{\mathbf{Q}}^{-1} \mathbf{a}_0 \cdot \mathbf{z}^H \hat{\mathbf{Q}}^{-1} \mathbf{z}}. \quad (21)$$

The relation to the generalised cosine, which is defined as  $\cos \angle(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a}^H \mathbf{b}}{\|\mathbf{a}\| \cdot \|\mathbf{b}\|}$ , is obvious. The statistics are related by  $T_{GLRT} = T_{AMF} / (1 + \frac{1}{K} \mathbf{z}^H \hat{\mathbf{Q}}^{-1} \mathbf{z})$  and  $T_{ACE} = T_{AMF} / \mathbf{z}^H \hat{\mathbf{Q}}^{-1} \mathbf{z}$ .

### Interpretations of the statistics and consequences:

1. The AMF detector represents an estimate of the signal-to-noise ratio because  $T_{AMF} = \frac{|\mathbf{w}^H \mathbf{z}|^2}{\mathbf{w}^H \hat{\mathbf{Q}} \mathbf{w}}$ . It has thus a physically intuitive interpretation.
2. The AMF detector may be interpreted as a CFAR detector with CFAR threshold averaged over secondary (training) beam output data, because  $T_{AMF} < \eta$  is equivalent to

$$|\mathbf{w}^H \mathbf{z}|^2 < \eta \cdot \frac{1}{K} \sum_{k=1}^K |\mathbf{w}^H \mathbf{z}_k|^2 = \eta \cdot \mathbf{w}^H \left( \frac{1}{K} \sum_{k=1}^K \mathbf{z}_k \mathbf{z}_k^H \right) \mathbf{w}.$$

3. The ACE detector can be interpreted as an adaptive sidelobe blanking scheme, because  $T_{ACE} < \eta$  is equivalent to  $|\mathbf{w}^H \mathbf{z}|^2 < \eta \cdot \mathbf{w}^H \mathbf{a}_0 \cdot G(\mathbf{z})$ , where  $G(\mathbf{z}) = \mathbf{z}^H \underbrace{\hat{\mathbf{Q}}^{-1}}_{=\mathbf{L}\mathbf{L}^H} \mathbf{z} = \underbrace{\mathbf{L}^H \mathbf{z}}_{=\mathbf{z}_{preWh}} \|\mathbf{z}\|^2 = \sum_{i=1}^N |z_{preWh,i}|^2$  is an adaptive guard channel calculated from the incoherent sum of all pre-whitened channels. The quantity  $\mathbf{w}^H \mathbf{a}_0$  is a normalisation of the weight vector which may be chosen equal to 1 (for a look direction constraint weight).
4. Similarly, the GLRT may be interpreted as an adaptive sidelobe blanking detector with a guard channel with constant level added.

The detection performance of the GLRT has been derived in [11], of the AMF in [12], and for the ACE in [13]. A compact derivation of the detection and false alarm probabilities of these detectors is given in [14]. These results show that the GLRT has superior performance over AMF and ACE. However, the results are only valid for an antenna without tapering. The case of a tapered antenna has been shortly characterised in [16] (called the mismatch case there). We give here only the final formulas for the detection and false alarm probabilities of these detectors in the general and compact form given in [14], [16] for tapered antennas (or look direction mismatch). The derivation of these probabilities is quite challenging. For details the reader is referred to these references. The distribution of the statistics is

$$\begin{aligned} T_{GLRT}^0 &\sim F_{1,K-N+1}(\delta_\beta), \quad \text{with } T_{GLRT}^0 = T_{GLRT} / (K - T_{GLRT}) \\ T_{AMF}^0 &\sim F_{1,K-N+1}(\delta_\beta) / \beta_{K-N+2,N-1}(c), \quad \text{with } T_{AMF}^0 = T_{AMF} / K \\ T_{ACE}^0 &\sim F_{1,K-N+1}(\delta_\beta) / (1 - \beta_{K-N+2,N-1}(c)), \quad \text{with } T_{ACE}^0 = T_{ACE} / (1 - T_{ACE}) \end{aligned}, \quad (22)$$

where  $F_{r,s}(\delta)$ , resp.  $\beta_{r,s}(c)$ , denotes a complex non-centrally F-, resp. Beta-, distributed random variable with  $r, s$  degrees of freedom and non-centrality parameter  $\delta_\beta$ , resp.  $c$ . The non-centrality parameters are

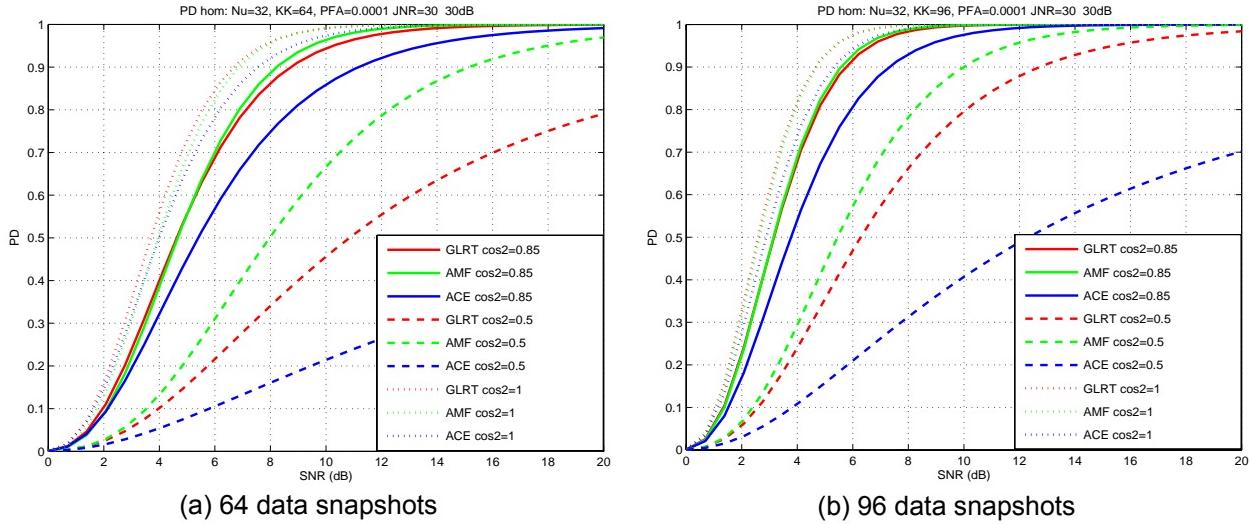
$$\begin{aligned} \delta_\beta^2 &= |S|^2 \beta_{K-N+2,N-1}(c) \cdot \tilde{\mathbf{m}}^H \mathbf{R}^{-1} \tilde{\mathbf{m}} \cos^2(\tilde{\mathbf{a}}_0, \tilde{\mathbf{m}}) \\ c^2 &= |S|^2 \cdot \tilde{\mathbf{m}}^H \mathbf{R}^{-1} \tilde{\mathbf{m}} \sin^2(\tilde{\mathbf{a}}_0, \tilde{\mathbf{m}}) \end{aligned}, \quad (23)$$

where  $|S|^2$  denotes the input signal power and  $\cos^2(\tilde{\mathbf{a}}_0, \tilde{\mathbf{m}}) = \frac{|\tilde{\mathbf{a}}_0^H \mathbf{Q}^{-1} \tilde{\mathbf{m}}|^2}{\tilde{\mathbf{a}}_0^H \mathbf{Q}^{-1} \tilde{\mathbf{a}} \cdot \tilde{\mathbf{m}}^H \mathbf{Q}^{-1} \tilde{\mathbf{m}}}$  and  $\sin^2(\tilde{\mathbf{a}}_0, \tilde{\mathbf{m}}) = 1 - \cos^2(\tilde{\mathbf{a}}_0, \tilde{\mathbf{m}})$ . For the case of perfect signal match (e.g. untapered antenna) with  $\tilde{\mathbf{a}}_0 = \tilde{\mathbf{m}}$  we have  $\cos^2(\tilde{\mathbf{a}}_0, \tilde{\mathbf{m}}) = 1$  and  $c = 0$ . The detection probability for a detection threshold  $\eta$  is now calculated by integrating these statistics  $T$  over their distribution  $P_T$ ,  $PD = \int_{\eta}^{\infty} T \cdot dP_T$ . The false alarm probability is obtained by inserting the distribution of the null hypothesis, i.e. the case when there is no target ( $|S|^2 = 0$ ), which means using the central distributions. The integrals are evaluated numerically by integrating the F-distribution for given  $\beta$  over the Beta-density.

The above formulas are valid for any data snapshot vector and any adaptive beamforming vector. This means that these formulas can be generalised to space-time adaptive processing (STAP).

Figure 13 shows the detection probability for the generic array with and without -35 dB Taylor tapering. We have applied a partial weighting at the array elements such that the noise power normalisation (or pre-whitening) as described in Sect. 4.5 is performed. The final taper weighting is applied at the subarrays. The false alarm probability was set to  $10^{-4}$ . Two jammers with 30 dB SNR at the elements at azimuth angles  $u = (-0.8, -0.4)$  corresponding  $(-53.1^\circ, -23.5^\circ)$  are present and adaptively suppressed with snapshot

numbers  $K = 2L$  and  $3L$  (number of subarrays  $L=32$ ). The tapering has a significant influence on the detection probability. For the generalized cosine we have already  $\cos^2(\mathbf{a}_0, \mathbf{m})=0.85$  (solid lines). In addition we have plotted PD curves for  $\cos^2(\mathbf{a}_0, \mathbf{m})=0.5$ , which corresponds to a target at half beamwidth off the look direction (dashed lines). In the untapered case the GLRT is best, but the AMF is very close to it (dotted lines). The GLRT is therefore considered as a sort of benchmark test. The ACE is very well suited to suppress sidelobe detections or any kind of mismatched signals.



**Figure 13: Detection probabilities for GLRT, AMF and ACE with and without tapering (homogeneous case).**

These analytical results are less suited to predict the detection performance of a real system. One would rather use them to study scenario and antenna dependencies. The detection probabilities of advanced adaptive weights, like subspace based or CAPS method, are too complicated to be determined analytically. These have to be evaluated by simulation. In addition, for a real radar other CFAR (constant false alarm rate) criteria may be introduced.

## 6.0 ANGLE ESTIMATION WITH ADAPTIVE ARRAYS

For adapted sum and difference beams the monopulse formula derived in Lecture 1 may produce large errors because the sum and difference beams can be severely distorted due to the nulls in the interference directions. This is in particular the case for jammers on the skirt of the main beam. However, the maximum likelihood estimation procedure described in Lecture 1 can be applied for an unknown interference covariance matrix, too. In the same way as in Lecture 1 one can derive a monopulse formula. This has been described in [17]. However, the ML procedure requires a special difference beam weight, which is the derivative of the sum beam weight. If we apply the monopulse formula to the subarray outputs, the true derivative cannot be applied. Also, often difference beams with low sidelobes are desired. An adaptive monopulse formula for arbitrary sum and difference beams has been derived in [18]. This formula has the form

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} - \begin{pmatrix} c_{xx} & c_{xy} \\ c_{yx} & c_{yy} \end{pmatrix} \begin{pmatrix} R_x - \mu_x \\ R_y - \mu_y \end{pmatrix}, \quad (24)$$

where  $\mathbf{C} = \begin{pmatrix} c_{xx} & c_{xy} \\ c_{yx} & c_{yy} \end{pmatrix}$  is a slope correction matrix and  $\boldsymbol{\mu} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}$  is a bias correction.  $R_x = \text{Re}\left\{\mathbf{d}_x^H \mathbf{z} / \mathbf{w}^H \mathbf{z}\right\}$

is the monopulse ratio formed with the measured difference and sum beam outputs  $D_x = \mathbf{d}_x^H \mathbf{z}$ ,  $S = \mathbf{w}^H \mathbf{z}$ , respectively, with arbitrary difference and sum beam weightings  $\mathbf{d}_x$ ,  $\mathbf{w}$  (and analogous for elevation estimation  $\mathbf{d}_y$ ). The bias correction is calculated as

$$\mu_\alpha = \operatorname{Re} \left\{ \frac{\mathbf{d}_\alpha^H \mathbf{a}_0}{\mathbf{w}^H \mathbf{a}_0} \right\}, \text{ for } \alpha = x, y \quad (25)$$

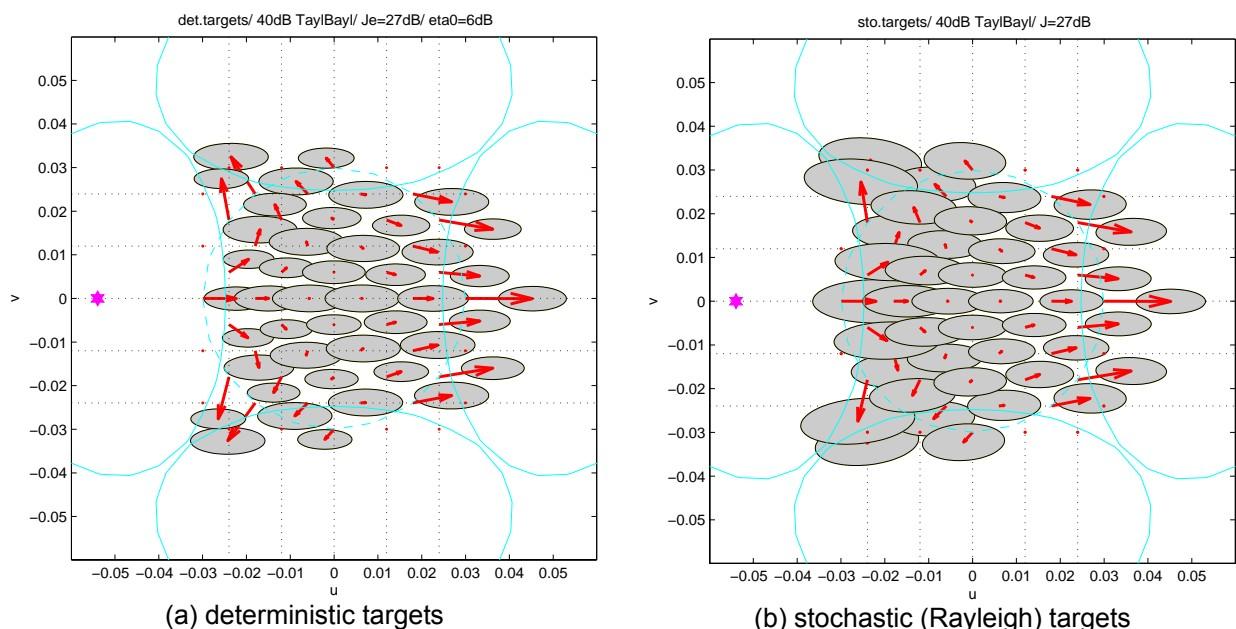
and the inverse slope correction matrix  $(c^{\alpha,h})_{\substack{a=x,y \\ h=u,v}} = \mathbf{C}^{-1}$  as

$$c^{\alpha,h} = \frac{\operatorname{Re} \left\{ \mathbf{d}_\alpha^H \mathbf{a}_{h,0} \mathbf{a}_0^H \mathbf{w} + \mathbf{d}_\alpha^H \mathbf{a}_0 \mathbf{a}_{h,0}^H \mathbf{w} \right\}}{\left| \mathbf{w}^H \mathbf{a}_0 \right|^2} - \mu_\alpha 2 \operatorname{Re} \left\{ \frac{\mathbf{w}^H \mathbf{a}_{h,0}}{\mathbf{w}^H \mathbf{a}_0} \right\} \quad (26)$$

for  $\alpha = x, y$ , and  $h = u, v$ .  $\mathbf{a}_0$  is the response of the array to a plane wave in the look direction and  $\mathbf{a}_{h,0}$  the derivative. In general these are fixed antenna determined quantities.

This monopulse formula is valid for any data snapshot vector and any adaptive beamforming vectors. In particular it can be applied to the subarray outputs or to a GSLC, which is a very special subarray configuration. Also the monopulse formula can be generalised to space-time adaptive processing. This has been formulated in [19]. Furthermore, the statistical distribution has been calculated in [19] for given weight vectors (conditional distributions).

Figure 14 shows the bias and standard deviation of the monopulse estimates for a jammer at one beamwidth off the antenna look direction (the jammer position is indicated by the asterisk). The 3 dB contours of the unadapted sum and difference beams are also indicated. Note that the jammer is nearly on the maximum of the difference beam which means that the difference pattern will be heavily distorted due to the adaptive null. With the performed slope and bias correction a reasonable angle estimate is obtained.



**Figure 14: Bias and standard deviation of monopulse estimates with adaptive beamforming for deterministic and stochastic targets, one jammer at 1 bw off antenna look direction.**

## 7.0 SUPERRESOLUTION METHODS

The resolution limit for classical beamforming is the 3 dB beamwidth of the sum beam. This limit is due to a signal processing that assumes only one target. With an array we have given spatial samples of the impinging wavefronts and one may assume a multi target model. This opens the possibility for enhanced resolution. Methods that yield a resolution beyond the conventional 3 dB beamwidth are called super-resolution methods. These methods have been discussed since decades, and textbooks on this topic are available, e.g. [20]. Of the many methods presented we consider here only those that can be applied to planar irregular spatial samples (which means also to irregular subarrays). We list below only some of the most important methods.

### Capon's method

$$S_C(\mathbf{u}) = \left( \mathbf{a}(\mathbf{u})^H \hat{\mathbf{R}}_{ML}^{-1} \mathbf{a}(\mathbf{u}) \right)^{-1} \text{ with } \hat{\mathbf{R}}_{ML} = \frac{1}{K} \sum_{k=1}^K \mathbf{z}_k \mathbf{z}_k^H$$

### MUSIC method (Multiple Signal Classification)

$$S_{MUSIC}(\mathbf{u}) = \left( \mathbf{a}(\mathbf{u})^H \mathbf{P}^\perp \mathbf{a}(\mathbf{u}) \right)^{-1} \text{ with } \mathbf{P}^\perp = \mathbf{I} - \mathbf{X} \mathbf{X}^H, \text{ and } \mathbf{X} \text{ spanning the dominant subspace}$$

### Deterministic ML method

$$F_{det}(\boldsymbol{\theta}) = \text{tr}(\mathbf{P}_A^\perp \hat{\mathbf{R}}_{ML}) \text{ with } \mathbf{P}_A^\perp = \mathbf{I} - \mathbf{A} \left( \mathbf{A}^H \mathbf{A} \right)^{-1} \mathbf{A}^H \text{ and } \mathbf{A}(\boldsymbol{\theta}) = (\mathbf{a}(\mathbf{u}_1), \dots, \mathbf{a}(\mathbf{u}_M))$$

### Stochastic ML method

$$F_{sto}(\boldsymbol{\theta}) = \log \det(\mathbf{R}(\boldsymbol{\theta})) + \text{tr}(\mathbf{R}(\boldsymbol{\theta})^{-1} \hat{\mathbf{R}}_{ML})$$

where  $\mathbf{R}(\boldsymbol{\theta})$  denotes the completely parameterised covariance matrix. One can give a formulation with the unknown directions as the only parameters

$$F_{sto}(\boldsymbol{\theta}) = \det \left\{ \mathbf{A}(\boldsymbol{\theta}) \mathbf{B}(\boldsymbol{\theta}) \mathbf{A}^H(\boldsymbol{\theta}) + \sigma^2(\boldsymbol{\theta}) \mathbf{I} \right\} \text{ with}$$

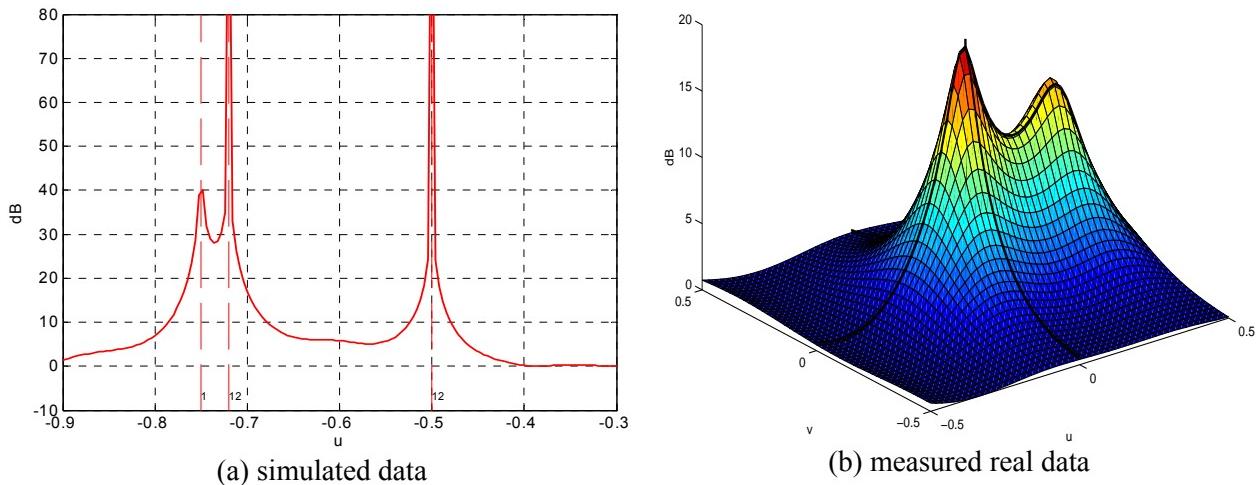
$$\sigma^2(\boldsymbol{\theta}) = \frac{1}{N-M} \text{tr} \left\{ \mathbf{P}_A^\perp \hat{\mathbf{R}}_{ML} \right\}, \quad \mathbf{B}(\boldsymbol{\theta}) = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \left( \hat{\mathbf{R}}_{ML} - \sigma^2(\boldsymbol{\theta}) \mathbf{I} \right) \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1} \text{ for } \mathbf{A} = \mathbf{A}(\boldsymbol{\theta}).$$

The deterministic ML method has some intuitive interpretations:

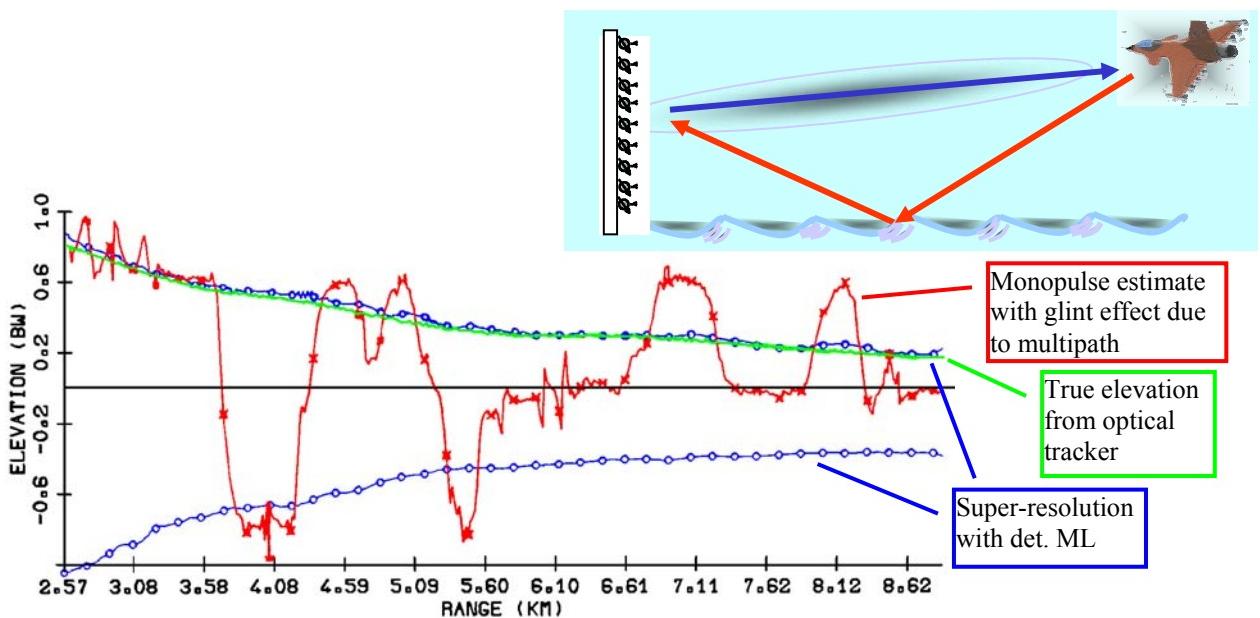
1.  $F_{det}(\boldsymbol{\theta}) = \frac{1}{K} \sum_{k=1}^K \mathbf{z}_k^H \mathbf{P}_A^\perp \mathbf{z}_k = \frac{1}{K} \sum_{k=1}^K \left\| \mathbf{P}_A^\perp \mathbf{z}_k \right\|^2 = \frac{1}{K} \sum_{k=1}^K \left\| \mathbf{z}_k - \underbrace{\mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{z}_k}_{=\hat{\mathbf{b}}} \right\|^2, \text{ which means that the mean squared residual error after signal extraction is minimised}$
2.  $F_{det}(\boldsymbol{\theta}) = C - \sum_{k=1}^K \mathbf{z}_k^H \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{z}_k, \text{ which can be interpreted as maximising a set of decoupled sum beams } (\mathbf{a}^H(\mathbf{u}_1) \mathbf{z}_k, \dots, \mathbf{a}^H(\mathbf{u}_M) \mathbf{z}_k).$
3.  $F_{det}(\boldsymbol{\theta}) = C - \mathbf{a}_{null}^H \hat{\mathbf{R}} \mathbf{a}_{null} / \mathbf{a}_{null}^H \mathbf{a}_{null} \text{ with } \mathbf{a}_{null} = \mathbf{P}_A^\perp \mathbf{a}, \text{ where we have partitioned the matrix of steering vectors into } \mathbf{A} = (\mathbf{a}, \hat{\mathbf{A}}). \text{ This property is valid due to the projection decomposition lemma which says}$

that for any partitioning  $\mathbf{A} = (\mathbf{F}, \mathbf{G})$  we can write  $\mathbf{P}_A^\perp = \mathbf{P}_G^\perp - \mathbf{P}_G^\perp \mathbf{F} (\mathbf{F}^H \mathbf{P}_G^\perp \mathbf{F})^{-1} \mathbf{F}^H \mathbf{P}_G^\perp$ . If we keep the directions in  $\hat{\mathbf{A}}$  fixed, this relation says that we have to maximise the scan pattern while the sources in the directions of  $\hat{\mathbf{A}}$  are deterministically nulled. One can now perform the multi-dimensional maximisation by alternating 1-dimensional maximisations and keeping the remaining directions fixed. This is the basis of the alternating projection method or IMP (Incremental Multi-Parameter) method, [20], p. 105.

The performance of the MUSIC method is illustrated in Figure 15 and for the deterministic ML method in Figure 16.



**Figure 15: MUSIC spectra.**



**Figure 16: Super-resolution of multipath propagation over sea with deterministic ML method (real data from vertical linear array with 32 elements, scenario illustrated above).**

The problems of these methods are described in [21], [22]. Main problems are numerical effort for finding the  $M$  maxima (one  $M$ -dimensional optimisation or  $M$  1-dimensional optimisations). To mitigate this

problem a stochastic approximation algorithm or the IMP method has been proposed for the deterministic ML method. Another problem is the exact knowledge of the signal model for all possible directions. This is sometimes called the array manifold. This is mainly an antenna accuracy or calibration problem. For an array with subarrays it is possible to perform super-resolution only with the subarray outputs. If one is only interested in super-resolution in the vicinity of the look direction, then one can use a simplified array manifold model based on the subarray centres, called the Direct Uniform Manifold model (DUM), which requires less calibration effort, [23]. Another problem is the determination of the number of sources. This can be achieved by the information theoretic methods presented in (15) or by the threshold test of (17).

## REFERENCES

- [1] L.S. Reed, J.D. Mallett, L.E. Brennan: Rapid convergence rate in adaptive arrays. IEEE Trans. AES 10 (6), 1974, pp. 853-863.
- [2] U. Nickel: Adaptive Beamforming for Phased Array Radars. Proc. Int. Radar Symposium IRS'98 (Munich, Sept. 1998), DGON and VDE/ITG, pp. 897-906.(Reprint also in Ortung und Navigation 2/98, pp. 99-108).
- [3] C.H. Gierull: Fast and effective method for low-rank interference suppression in presence of channel errors. IEE Electronic Letters 34 (6), 1998, pp. 518-520.
- [4] G.M. Herbert: A new projection based algorithm for low sidelobe pattern synthesis in adaptive arrays. IEE Int. Conf. Radar 97, 14-17 Oct. 1997, IEE Publ. Nr. 449, pp. 396-400.
- [5] U. Nickel: Determination of the Dimension of the Signal Subspace for Small Sample Size. Proc. IASTED Int. Conf. on Signal Processing and Comm., IASTED/Acta Press 1998, pp. 119-122.
- [6] G.H. Golub, C.F. van Loan: Matrix Computations. Johns Hopkins Univ. Press, 1996 (3<sup>rd</sup> Ed).
- [7] U. Nickel: On the Influence of Channel Errors on Array Signal Processing Methods. Int. J. Electronics and Comm. (AEÜ) 47 (4), 1993, pp. 209-219.
- [8] A.B. Gershman, U. Nickel, J.F. Böhme: Adaptive Beamforming Algorithms with Robustness against Jammer Motion. IEEE Trans. SP-45 (7) 1997, pp. 1878-1885.
- [9] SD Hayward: Adaptive Beamforming for Rapidly Moving Arrays, CIE Int. Conf. on Radar ICR96, Beijing, IEEE Cat.Nr. 96TH8117.
- [10] E.J. Kelly: An adaptive detection algorithm. IEEE Trans. AES 22 (1), 1986, pp. 115-127.
- [11] E.J. Kelly: Performance of an adaptive detection algorithm; rejection of unwanted signals. IEEE Trans. AES 25 (2), 1989, pp. 122-133.
- [12] F.C. Robey, D.R. Fuhrmann, E.J. Kelly, R. Nitzberg: A CFAR adaptive matched filter detector. IEEE Trans. AES 28 (1) 1992, pp. 208-216.
- [13] S. Kraut, L.L. Scharf, L.T. McWhorter: Adaptive subspace detectors. IEEE Trans. SP 49 (1) 2001, pp. 1-16.
- [14] C.D. Richmond: Performance of the adaptive sidelobe blanker detection algorithm in homogeneous environments. IEEE Trans. SP 48 (5) 2000., pp. 1235-1247.

- [15] C.D. Richmond: Performance of a class of adaptive detection algorithms in non-homogeneous environments. IEEE Trans. SP 48 (5) 2000., pp. 1248-1262.
- [16] C.D. Richmond: Theoretical performance of a class of space-time adaptive detection and training strategies for airborne radar. Asilomar Conf. on Signals, Systems and Computers, Pacific Grove, CA, 1998, IEEE Cat.Nr. 0-7803-5148-7/98, pp. 1327-1331.
- [17] U. Nickel: Monopulse estimation with adaptive arrays. IEE Proc.-F 140 (5), 1993, pp. 303-308.
- [18] U. Nickel: Monopulse estimation with subarray-adaptive arrays and arbitrary sum and difference beams. IEE Proc. Radar Sonar Navig. 143 (4), 1996, pp. 232-238.
- [19] U. Nickel: Performance analysis of space-time adaptive monopulse. Signal Processing 84 (9), 2004, pp. 1561-1579.
- [20] S. Haykin (ed): Advances in Spectrum Analysis and Array Processing, Vol. II. Prentice Hall 1991.
- [21] U. Nickel: Radar target parameter estimation with antenna arrays, in S.Haykin, J.Litva, T.J.Shepherd (eds): Radar Array Processing, Springer Verlag 1993, pp. 47-98.
- [22] U. Nickel: Aspects of implementing super-resolution methods into phased array radar. Int. J. of Electron. Comm. (AEÜ), 53 (6), 1999, pp. 315-323.
- [23] U. Nickel: Spotlight MUSIC: Super-resolution with subarrays with low calibration effort. IEE Proc. Radar, Sonar, Navigation Vol. 149 (4), Aug. 2002, pp. 166-173.